

Some comments about my RDA participation

André Schaaff

CDS, Centre de Données astronomiques de Strasbourg

IVOA Cape Town, DCP session





□ Outlines

WG, IG, BoF participation in Tokyo
Focus on topics related to DCP

□ IG PID

- **Aim:** Synchronize identifier-related efforts, coordinate activities, solve PID-related issues from the participant communities, PID use cases, collaboration with the other WG
- PID interoperability: Crossref, Datacite, ORCID (over 2 million researchers have an ORCID id), ORCID use supported by the publishers
- Organization IDs: on-going project (like in Jisc CASRAI-UK) working around the identification of organisations (currently no single (global-interoperable-unambiguous-unique) institutional identifier)
- Connections between datasets and publications, no common standard, Data Literature Interlinking and OpenAire as reference implementation
- Example of IGSN, International Geo Sample Number, a unique alphanumeric code (Central catalog, descriptive metadata supports, etc.) that uniquely identifies samples from the natural environment, recommended by the publishers

Credits : BigHadoop

□ IG Repository Platforms for Research Data

- **Aim:** gather and analyse research data use cases in the context of repository platform requirements, influence the development of repository software and related services to better serve the use cases of the research data community
- List of functional requirements, ...

Credits : BigHadoop

RDA Repository Platforms for Research Data Interest Group

Use Case: Use Case Name

Author(s):

This is a use case description of the "Repository Platforms for Research Data" IG. While points 1, 2 and 3 aim at a general description/overview of the use case, point 4 is meant to list the requirements.

Please, save the file using the name scheme: UseCaseName_UseCase_RepoPlat.docx

1. Scientific Motivation and Outcomes

Provide a short summary of the scientific or technical motivation for the use case. What would be the best possible outcome and why?

2. Functional Description

If possible, give one diagram/picture that indicates the overall structure/architecture you have/envision. Describe in simple words the functioning of the use case/system.

3. Achieved Results

Describe results (if applicable) that have been achieved compared to the original motivation. What requirements could not be fulfilled and how did this influence the outcomes

4. Requirements

Describe the requirements, their motivation from your use case and how you rate their importance. The descriptions don't have to be comprehensive.

Requirement	Description	Motivation from Use Case	Importance (1 - very important to 5 - not at all important)

□ BoF on Research Data Repository Interoperability

- **Aim:** establish standards for interoperability between different repository platforms
- It includes the work on a generic API and import/export formats

Credits : BigHadoop

□ WG Research Data Collections

- **Aim:** examine existing models for identifying and managing collections of data
- Collection models: descriptions and use cases
- Propose an API specification (to expose the hierarchical levels) and a demonstrator (to show the added-value)
- PID

Credits : BigHadoop

□ Big Data: Considerations in Applying Big Data Technologies

- **Aim:** develop community based recommendations on feasible data analytics approaches to address scientific community needs of utilizing large quantities of data.
- Not in the frame of DCP but possible links in GWS through current work around Hadoop / Spark

Credits : BigHadoop

□ BoF on Text and Data Mining: Defining the Challenges and Actions

- **Aim:** produce a list of technical, legal, policy and challenges that content providers faced in making their data available for text and data mining
- **Audience:** for data provider who are interested in making their data available for text and data mining

Credits : BigHadoop

□ BoF on IG-VRE (Virtual Research Environment)

- **Aim:** architectural principles, technologies and best practice in the design and development and of VREs
- Comment...

Credits : BigHadoop

□ IG Federated Identify Management

- Not in the frame of DCP, but as information
- “The FIM interest group (FIMig) is an international cross-domain interest group to work on all issues related to the use FIM for the implementation of AAls (Authentication and Authorization Infrastructures) in research infrastructures. “
- But no session in Tokyo

Credits : BigHadoop